Computer-based intensity measurement assists pathologists in scoring PTEN immunohistochemistry - Clinical associations in NSCLC patients of the ETOP Lungscape cohort

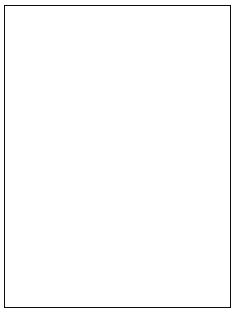
Rulle, Undine ; Tsourti, Zoi ; Casanova, Ruben ; Deml, Karl-Friedrich ; Verbeken, Eric ; Thunnissen, Erik ; Warth, Arne ; Cheney, Richard ; Sejda, Aleksandra ; Speel, Ernst Jan ; Madsen, Line Bille ; Nonaka, Daisuke ; Navarro, Atilio ; Sansano, Irene ; Marchetti, Antonio ; Finn, Stephen P ; Monkhorst, Kim ; Kerr, Keith M ; Haberecker, Martina ; Wu, Chengguang ; Zygoura, Panagiota ; Kammler, Roswitha ; Geiger, Thomas ; Gendreau, Steven ; Schulze, Katja ; Vrugt, Bart ; Wild, Peter ; Moch, Holger ; Weder, Walter ; Ciftlik, Ata Tuna ; Dafni, Urania ; Peters, Solange ; Bubendorf, Lukas ; Stahel, Rolf A ; Soltermann, Alex

Abstract: INTRODUCTION PTEN loss is frequently observed in NSCLC and associated with both PI3K activation and tumoral immunosuppression. PTEN immunohistochemistry is a valuable read-out, but lacks standardized staining protocol and cut-off value. METHODS Following an external quality assessment (EQA) using SP218, 138G6 and 6H2.1 anti-PTEN antibodies, scored on webbook and tissue microarray, the ETOP cohort samples (n=2245 NSCLC patients, 8980 TMA cores) were stained with SP218. All cores were H-scored by pathologists and by computerized pixel-based intensity measurements calibrated by pathologists. RESULTS All 3 antibodies differentiated 6 PTEN+ versus 6 PTEN- cases on EQA. For 138G6 and SP218, high sensitivity and specificity was found for all H-score threshold values including prospectively defined 0, calculated 8 (pathologists) and calculated 5 (computer). High concordance among pathologists in setting computer-based intensities and between pathologists and computer in H-scoring was observed. Due to over-integration of the human eye, pixel-based computer H-scores were overall 54% lower. For all cut-off values, PTEN- was associated with smoking history, squamous cell histology and higher tumor stage ($p<0.001$). In adenocarcinomas, PTEN- was associated with poor survival. CONCLUSION Calibration of immunoreactivity intensities by pathologists, following computerized H-score measurements has the potential to improve reproducibility and homogeneity of biomarker detection regarding epitope validation in multi-centre studies.

DOI: https://doi.org/10.1016/j.jtho.2018.08.2034

Originally published at:
Rulle, Undine; Tsourti, Zoi; Casanova, Ruben; Deml, Karl-Friedrich; Verbeken, Eric; Thunnissen, Erik; Warth, Arne; Cheney, Richard; Sejda, Aleksandra; Speel, Ernst Jan; Madsen, Line Bille; Nonaka, Daisuke; Navarro, Atilio; Sansano, Irene; Marchetti, Antonio; Finn, Stephen P; Monkhorst, Kim; Kerr, Keith M; Haberecker, Martina; Wu, Chengguang; Zygoura, Panagiota; Kammler, Roswitha; Geiger, Thomas; Gendreau, Steven; Schulze, Katja; Vrugt, Bart; Wild, Peter; Moch, Holger; Weder, Walter; Ciftlik, Ata Tuna; Dafni, Urania; Peters, Solange; Bubendorf, Lukas; Stahel, Rolf A; Soltermann, Alex (2018). Computer-based intensity measurement assists pathologists in scoring PTEN immunohistochemistry - Clinical associations in NSCLC patients of the ETOP Lungscape cohort. Journal of Thoracic Oncology, 13(12):1851-1863.
DOI: https://doi.org/10.1016/j.jtho.2018.08.2034

Accepted Manuscript



Computer-based intensity measurement assists pathologists in scoring PTEN immunohistochemistry - Clinical associations in NSCLC patients of the ETOP Lungscape cohort

Undine Rulle, Zoi Tsourti, Ruben Casanova, Karl-Friedrich Deml, Eric Verbeken, Erik Thunnissen, Arne Warth, Richard Cheney, Aleksandra Sejda, Ernst Jan Speel, Line Bille Madsen, Daisuke Nonaka, Atilio Navarro, Irene Sansano, Antonio Marchetti, Stephen P. Finn, Kim Monkhorst, Keith M. Kerr, Martina Haberecker, Chengguang Wu, Panagiota Zygoura, Roswitha Kammler, Thomas Geiger, Steven Gendreau, Katja Schulze, Bart Vrugt, Peter Wild, Holger Moch, Walter Weder, Ata Tuna Ciftlik, Urania Dafni, Solange Peters, Lukas Bubendorf, Rolf A. Stahel, Alex Soltermann

PII: S1556-0864(18)33053-3

DOI: 10.1016/j.jtho.2018.08.2034

Reference: JTHO 1130

To appear in: *Journal of Thoracic Oncology*

Received Date: 19 February 2018

Revised Date: 16 July 2018

Accepted Date: 2 August 2018

Please cite this article as: Rulle U, Tsourti Z, Casanova R, Deml K-F, Verbeken E, Thunnissen E, Warth A, Cheney R, Sejda A, Speel EJ, Madsen LB, Nonaka D, Navarro A, Sansano I, Marchetti A, Finn SP, Monkhorst K, Kerr KM, Haberecker M, Wu C, Zygoura P, Kammler R, Geiger T, Gendreau S, Schulze K, Vrugt B, Wild P, Moch H, Weder W, Ciftlik AT, Dafni U, Peters S, Bubendorf L, Stahel RA, Soltermann A, Computer-based intensity measurement assists pathologists in scoring PTEN immunohistochemistry - Clinical associations in NSCLC patients of the ETOP Lungscape cohort, *Journal of Thoracic Oncology* (2018), doi: https://doi.org/10.1016/j.jtho.2018.08.2034.

Computer-based intensity measurement assists pathologists in scoring PTEN immunohistochemistry - Clinical associations in NSCLC patients of the ETOP Lungscape cohort

Authors

Undine Rulle[a*], Zoi Tsourti[b*], Ruben Casanova[a], Karl-Friedrich Deml[c], Eric Verbeken[d], Erik Thunnissen[f], Arne Warth[g], Richard Cheney[h], Aleksandra Sejda[i], Ernst Jan Speel[j], Line Bille Madsen[k], Daisuke Nonaka[l], Atilio Navarro[m], Irene Sansano[n], Antonio Marchetti[o], Stephen P. Finn[p], Kim Monkhorst[q], Keith M. Kerr[r], Martina Haberecker[a], Chengguang Wu[a], Panagiota Zygoura[b], Roswitha Kammler[s], Thomas Geiger[s], Steven Gendreau[t], Katja Schulze[t], Bart Vrugt[a], Peter Wild[a], Holger Moch[a], Walter Weder[u], Ata Tuna Ciftlik[v], Urania Dafni[w], Solange Peters[x], Lukas Bubendorf[y], Rolf A. Stahel[z], Alex Soltermann[a].

* UR and ZT share the first authorship

Affiliations

a) Institute of Pathology and Molecular Pathology, University Hospital Zurich, 8091 Zurich, Switzerland;

b) Frontier Science Foundation Hellas, 15773 Athens, Greece;

c) Institute of Pathology, Cantonal Hospital Münsterlingen, 8596 Münsterlingen, Switzerland;

d) Institute of Pathology, University Hospitals Leuven, 3000 Leuven, Belgium;

f) Institute of Pathology, VU University Medical Centre, 1081 Amsterdam, Netherlands;

g) Institute of Pathology, University Hospital Heidelberg, 69120 Heidelberg, Germany;

h) Department of Pathology and Anatomical Science, State University of New York at Buffalo, Buffalo, New York, United States of America, NY 14263;

i) Institute of Pathology, Medical University of Gdansk, 80-210 Gdansk, Poland;

j) Institute of Pathology, Maastricht University Medical Centre, 6211 Maastricht, Netherlands;

k) Institute of Pathology, Aarhus University Hospital, 8200 Aarhus, Denmark;

l) Institute of Pathology, Christie NHS Foundation Trust, M20 4BX, Manchester, United Kingdom;

m) Institute of Pathology, University General Hospital of Valencia, 46014 Valencia, Spain;

n) Institute of Pathology, Vall d`Hebron University Hospital, 08035 Barcelona, Spain;

o) Institute of Pathology, Ospedale Clinicizzato, 66100 Chieti, Italy;

p) Departments of Histopathology and Cancer Molecular Diagnostics, St James's Hospital and Trinity College, 152-160 Dublin, Ireland;

q) Institute of Pathology, Netherlands Cancer Institute, 1066 Amsterdam, Netherlands;

r) Institute of Pathology, Aberdeen Royal Infirmary, AB25 2ZN Aberdeen, United Kingdom;

s) Translational Research Coordination - European Thoracic Oncology Platform, 3008 Bern, Switzerland;

t) Genentech Inc., San Francisco, California, United States of America, CA 94080;

u) Division of Thoracic Surgery, University Hospital Zurich, 8091 Zurich, Switzerland;

v) Lunaphore Technologies SA, EPFL Innovation Park, 1015 Lausanne, Switzerland;

w) Frontier Science Foundation Hellas & University of Athens, 15773 Athens, Greece;

x) Department of Oncology, University Hospital Lausanne, 1011 Lausanne, Switzerland;

y) Institute of Pathology, University Hospital Basel, 4031 Basel, Switzerland;

z) Clinic and Policlinic of Oncology, University Hospital Zurich, 8091 Zurich, Switzerland.

Corresponding author

Prof. Alex Soltermann MD, Institute of Pathology and Molecular Pathology, University Hospital Zurich, Schmelzbergstrasse 12, CH-8091, Zurich, Switzerland. Tel: ++41 44 255 23 19; Fax: ++41 44 255 45 52; alex.soltermann@usz.ch

Competing interests

Steven Gendreau and Katja Schulze are employees of Genentech Inc.

Abstract

Introduction: PTEN loss is frequently observed in NSCLC and associated with both PI3K activation and tumoral immunosuppression. PTEN immunohistochemistry is a valuable read-out, but lacks standardized staining protocol and cut-off value.

Methods: Following an external quality assessment (EQA) using SP218, 138G6 and 6H2.1 anti-PTEN antibodies, scored on webbook and tissue microarray, the ETOP cohort samples (n=2245 NSCLC patients, 8980 TMA cores) were stained with SP218. All cores were H-scored by pathologists and by computerized pixel-based intensity measurements calibrated by pathologists.

Results: All 3 antibodies differentiated 6 PTEN+ versus 6 PTEN- cases on EQA. For 138G6 and SP218, high sensitivity and specificity was found for all H-score threshold values including prospectively defined 0, calculated 8 (pathologists) and calculated 5 (computer). High concordance among pathologists in setting computer-based intensities and between pathologists and computer in H-scoring was observed. Due to over-integration of the human eye, pixel-based computer H-scores were overall 54% lower. For all cut-off values, PTEN- was associated with smoking history, squamous cell histology and higher tumor stage ($p<0.001$). In adenocarcinomas, PTEN- was associated with poor survival.

Conclusion: Calibration of immunoreactivity intensities by pathologists, following computerized H-score measurements has the potential to improve reproducibility and homogeneity of biomarker detection regarding epitope validation in multi-centre studies.

Keywords

PTEN; NSCLC; external quality assessment; computer-based intensity measurement; immunohistochemistry.

Introduction

Phosphatase and tensin homolog (PTEN) is a major tumor suppressor with pleiotropic functions on cell survival, proliferation and chromosomal integrity [1, 2]. Cytosolic loss of function (PTEN-) leads to hyperactivation of the PI3K/AKT/mTOR signaling pathway, an event which is common to non-small cell lung carcinoma (NSCLC) [3, 4]. Loss of nuclear PTEN leads to an unstable genome with increased mutational burden, sensitizing cells to DNA-targeting drugs such as cisplatin [5, 6]. There is no validated clinically effective special treatment for PTEN- cancers yet, but substantial research and drug development for the PTEN/PI3K axis is ongoing, in particular PI3K inhibitors, for which PTEN- is a surrogate marker. There are phase I and II clinical trials with PI3K inhibitors in solid tumors and NSCLC, respectively [7, 8] as well as in castration-resistant prostate carcinoma requiring assessment of PTEN status [9]. PTEN is not only an important predictive biomarker for PI3K inhibition, but it has recently been reported that PTEN- might be a mechanism of resistance to cancer immunotherapy, e.g. in metastatic uterine leiomyosarcoma [10]. In mouse melanoma cells PTEN- promoted resistance to T-cell mediated immunotherapy such as anti-PD1/PD-L1, by decreasing T cell infiltration and expansion in tumors [11]. Such loss causes an immune-suppressive microenvironment by activation of T-regs [12] and inhibition of natural killer cells [13]. In lung squamous cell carcinomas (LSCC) loss of LKB1 and PTEN led to elevated PD-L1 expression [14]. However, in high grade lung neuroendocrine carcinoma, no significant correlation among PTEN loss, immune cell infiltration, and PD-L1 expression on either tumor cells or immune cells was found [15]. Finally, there are different miRNAs (mRNA-21, miRNA-92b, mRNA-21, mir-26b, miRNA-181a) regulating PTEN expression, thereby affecting cell growth, migration and resistance/sensitivity to platinum-based chemotherapies [16].

The poor prognostic value of PTEN- has been described in various cancers [17-19]. In NSCLC, two studies in 2012 showed that protein loss as measured by immunohistochemistry (IHC) occurs in up to 21% [20] or up to 59% [21] of LSCC. For lung adenocarcinomas (LADC) the

frequencies were 4% and 34%, respectively. Both studies used the 138G6 antibody and PTEN- was defined as absence of any immunoreactivity. An earlier study in 2005 found a complete loss or reduced PTEN protein expression in 74% of NSCLC, whereby reduced expression was defined as positive staining of any intensity in less than 50% of the tumor cells, using the 6H2.1 antibody [22]. In high grade lung neuroendocrine carcinoma, complete loss of PTEN protein was found in 9.5%, using the 28H6 antibody [15].

Clinical assessment of PTEN status is therefore important, but lacks standards. Determination of protein loss by IHC is considered the best approach, as it integrates various regulatory networks acting on the enzyme. IHC is assumed to be superior to sequencing since more PTEN- cases were detected by IHC than by sequencing in endometrial carcinoma [23]. In particular, PTEN protein loss was detected by IHC in 44% of cases classified as PTEN wild type by sequencing. In high grade lung neuroendocrine carcinomas, PTEN IHC expression had no correlation with PTEN mutation status assessed by genomic analysis [15]. Given that PTEN is a tumor suppressor, protein loss rather than overexpression is pathologically and clinically relevant. Therefore, the low level expression range close to H-score 0 is crucial to evaluate a potential PTEN-.

Mindful of the discrepant results for the 138G6 antibody, we first, evaluated different staining protocols for PTEN IHC across European Thoracic Oncology Platform (ETOP) laboratories and determined an optimal H-score threshold value for PTEN-. The SP218 clone was used alongside the established 138G6 and 6H2.1 antibodies. Pathologists' H-scores were compared with a novel approach of objective computerized pixel-based intensity measurement calibrated by pathologists. We investigated the prevalence of PTEN- and its correlation with clinic-pathologic data in the ETOP Lungscape cohort of 2245 resected NSCLC patients.

Material and methods

External quality assessment of PTEN immunohistochemistry

An external quality assessment (EQA) was performed using 3 different anti-PTEN antibodies (rabbit monoclonal SP218 antibody (Spring Bioscience, Cat. No. M5180), rabbit monoclonal 138G6 antibody (Cell Signaling Technologies, Cat No. 9559) and mouse monoclonal 6H2.1 antibody (DAKO, Cat. No. M3627)) on two automated IHC platforms (Ventana Benchmark® and Leica Bond-Max®). To avoid differences in preanalytical conditions, all ETOP labs used the Ventana Ultra automated platform with standardized reagents, in particular same batch of SP218 antibody distributed to all ETOP centres (Supplementary Table 1).

Overall, 12 cases were retrieved from our archive according to the PTEN IHC result indicated in the sign-out report using the 6H2.1 antibody. These 12 cases included 5 positive and 5 negative surgical specimens, respectively, and two cell lines, PTEN- PC3 [24] and PTEN+ H460. Firstly, whole sections were centrally stained and digitized on a webbook (Institute of Pathology Zurich) using all 3 antibodies. Secondly, these cases were assembled in tissue microarrays (TMA) and stained locally (each ETOP laboratory). Locally stained TMA sections were sent back to Zurich for computer analysis. To compare SP218 data with the 136G6 antibody, we performed IHC on the Zurich ETOP TMAs (n= 305 patients) with the latter antibody.

To train pathologists, a visual presentation was first given, including low expression cases of H-scores ≤50. IHC results were discussed at a follow-up meeting. Second, PTEN staining and H-scoring instructions were distributed. Pathologists received all EQA results from their colleagues and associated statistics. Scoring results were examined by U.R. and A.S. and discussed with pathologists. Third, permission for further staining and scoring of the entire ETOP cohort with the SP218 antibody was delivered.

ETOP NSCLC patient Lungscape cohort and website databases

The iBiobank has annotated comprehensive data from 16 ETOP sites on surgically resected 2245 patients with stage I to III NSCLC with at least 3 years of follow-up (median follow-up = 4.8 years) full clinical history [25, 26]. The iBiobank also contains data on *EGFR*, *BRAF*,

KRAS, *PIK3CA* alterations and ALK, MET IHC. Research was conducted according to each participating country ethics and regulatory requirements for use of patient's material in research. We additionally performed in-silico *PTEN* mRNA expression analysis from the Kaplan-Meier Plotter database (www.kmplot.com), consisting of 1145 lung cancer cases, including 673 LADC and 271 LSCC [27].

PTEN fluorescent in-situ hybridization

EQA TMA sections of 4 µm thickness were incubated with a dual color probe for cytoband 10q23 and region 10p11.1-q11.1 (LSI PTEN Spectrum-Orange and CEP10 Spectrum-Green, Vysis/Abbott Molecular, Baar, Switzerland). For each case, 100 non-overlapping nuclei were evaluated on a Zeiss Axioskop (Oberkochen, Germany). Z-stacks of 20 images with 0.5 µm step distance were merged.

H-scoring of PTEN IHC by pathologists and computer

ETOP Lungscape pathologists determined H-scores on webbook and TMA sections. The semi-quantitative score was obtained by the summation of the product of total cellular PTEN immunoreactivity intensity (0, 1, 2 or 3) with corresponding percentage of stained tumor epithelia (H-score range 0-300). Robust PTEN expression (score 2 or 3) in cancer-associated fibroblast (CAFs), endothelial cells and alveolar pneumocytes served as internal control.

Stained sections were digitalized using a NanoZoomer Digital Pathology scanner (Hamamatsu, Japan) at the maximum in-built magnification of 400x. Image analysis was done at 200x magnification on the virtual microscope software Leica SlidePath. Tagged image file format (TIFF) pictures of tumor epithelia only were analysed using Image-J2 software [28]. Pixel-based measurement of immunoreactivity was performed using the Lab* colour space. White areas lacking tissue were removed in the brightness channel, whereas blue counterstaining areas were removed from chrominance-a and -b. The resulting brown signal was thereafter located between red and yellow. The threshold between the background brown

signal and specific immunoreactivity intensity score 1 was averaged from 110 representative image frames laid on faintly stained glass (10 frames), unstained tissue areas of a microfluidic tissue processor-based IHC (50 frames) [29] and stainings without primary antibody (50 frames). This background setting was maintained for all further analyses. Subsequently, 3 Zurich pathologists and 10 ETOP pathologists set their individual thresholds of brown intensity scores 3, 2 and 1 for the SP218 antibody (Supplementary Video). The 138G6 antibody stainings were scored by 2 other Zurich investigators.

Computer H-scores were calculated in the same way by multiplying percentage of positive pixel area with respective intensity. For the one EQA whole section and the 4 EQA TMA cores of each of the 12 cases, 50 frames in total were put on tumor epithelia in a semi-automated manner. For the 97 TMAs of the 2245 patient ETOP cohort (n=8980 cores), 4 frames per case were used, covering largest possible areas of only vital tumor but not CAFs or necrosis. For the computerized analysis of the full cohort, the Zurich settings were applied on all images from externally stained TMAs in order to provide uniformity. The frame sizes were in the range of 100-500 x 100-500 pixels (1µm=2.17 pixels), generally being larger in the ETOP cohort than in the EQA whole sections or the EQA TMA cores (Supplementary Figure 1).

Statistical analysis

In the EQA dataset, Wilcoxon non-parametric test checked for significant differences between PTEN+ vs PTEN- cases. Sensitivity and specificity of H-score thresholds were calculated through ROC curves. In the ETOP cohort, the prevalence of PTEN- and confidence interval (95% CI) was compared between patients with different clinic-pathologic characteristics using the Fisher's exact and the Mantel-Haenszel tests. Associations between PTEN- and predictive markers were evaluated by 2-way tables and the Cochran-Mantel-Haenszel test, stratified by histology.

Clinical outcome included overall survival (OS, time from surgery to death from any cause), relapse-free survival (RFS, time from surgery to first relapse or death from any cause), and time-to-relapse (TTR, time from surgery to first relapse) [25, 26]. The effect of PTEN- on outcome was explored through Cox regression models, adjusted for a series of patient, tumor, surgery characteristics and for mutations. Final models with significant outcome prognostic factors were based on the backwards elimination method (Wald's $p \geq 0.10$). Hazard ratios (HRs) and Kaplan-Meier curves were used to illustrate observed differences in hazard.

In all exploratory analyses, results with two-sided $p \leq 0.05$ were considered significant. Analyses were carried out using SAS version 9.4 (SAS Institute, Cary, NC, USA) and R version 3.3.2 (R Foundation for Statistical Computing, Vienna, Austria).

Results

PTEN immunohistochemistry

Routine diagnostic protocol was used for 6H2.1 antibody, but for the SP218 and 138G6 anti-PTEN clones, we implemented novel protocols. To avoid differences in preanalytical conditions, all labs used the Ventana Ultra automated platform with standardized reagents and the same batch of SP218 antibody (Supplementary Table 1). Figure 1A/B shows the performance of the 3 antibodies in a PTEN- mucinous endometrial carcinoma and representative images for different H-scores. The methodology of computerized intensity measurement calibrated by pathologists is presented in Figure 2. Examples of EQA TMA staining quality among centers are presented in Supplementary Figure 2 for two cases.

External quality assessment

Table 1 summarizes IHC and FISH data from the 12-case TMA and whole sections webbook used for EQA. The webbook analysis yielded similar results for SP218 and 138G6 antibodies. 6H2.1 stained stronger than SP218 and 138G6 (Supplementary Table 2), as in Figure 1A for a single case (computer H-score 3 versus 0) and Figure 3A. All antibodies could separate

PTEN- from PTEN+ cases based on H-score average. All negative cases also demonstrated a genomic *PTEN* loss, ranging from 0.90 down to 0.02 CEP10/PTEN ratios. Correlation of H-scores among the antibodies was highly significant (p-values 0.002-0.003) with coefficients 0.78 (138G6/6H2.1), 0.78 (SP218/138G6) and 0.80 (SP218/6H2.1) (data not shown).

The trade-off between sensitivity and specificity, for the EQA TMA data, was more balanced for intermediate than high H-scores (ROC curves in Figures 3B). Analyses of these ROC curves showed that the optimal cut-off threshold of PTEN- is mean H-score <5 for computer and <8 for pathologists (values shown in Figures 3B). In order to address intratumoral heterogeneity, we calculated the variation coefficient among the H-scores of the 50 randomly selected frames of the PTEN+ webbook whole sections. A value of 0.38 was obtained for the SP218 antibody, indicating little variation.

ETOP NSCLC Lungscape cohort

As the two antibodies SP218 and 138G6 had similar performance in the EQA TMA, investigation was performed to assess if they are also comparable in a bigger sample size. The Zurich part of Lungscape (n=305) was therefore investigated by both SP218 and 138G6. Results of both pathologists' and computer scorings showed a strong correlation between these antibodies (Spearman correlation coefficient between the two antibodies: 0.73 for computer scores and 0.93 for pathologists scores; p<0.001 in both cases). SP218 was chosen for further analysis of the full ETOP Lungscape cohort.

In this cohort with available PTEN information, the median patient age is 66.4 years, 66% are males and 34% females. 53.9% of patients are former, 31.5% current and 10.6% non-smokers. Stage distribution is as follows: IA 22.4%, IB 25.9%, IIA 16.8%, IIB 12.2%, IIIA 20.8% and IIIB 1.9%. 48.7% of tumors are LADC, 43.8% LSCC, 4.4% large cell carcinoma (LCC) and 3.2% other histology. Median tumour size is 3.5cm. Available 5-year overall survival (OS) is 53.6% (95% CI:51.3%-55.9%), 5-year relapse free survival (RFS) is 46.5% (95% CI: 44.3%-48.8%), and 5-year time to relapse (TTR) is 56.7% (95% CI: 54.4%-59.0%).

Correlations of PTEN immunoreactivity with clinic-pathologic parameters are summarized in Table 2. PTEN-, using defined cut-offs for SP218, was detected in 1097/2245 patients (48.9%, based on the pathologist cut-off=8) or 1304/2208 (59.1%, based on the computer cut-off=5). In addition, when the H-score 0 cut-off was tested for the pathologists, 981/2245 patients (43.7%) were PTEN-. For all 3 thresholds PTEN- correlated with former and current smokers, LSCC histology, higher stage and larger tumor size. PTEN- according to pathologists' scores was more common to patients which obtained adjuvant chemotherapy as well as radiotherapy. Correlation with other biomarkers is presented in Supplementary Table 3. In brief PTEN- correlated with ALK and MET.

Overall (n=2245), the median OS was 67.9 months (95% CI: 62.0, 74.3), the median RFS was 50.8 months (95% CI: 45.8, 56.4) and the median TTR was 103 months (95% CI: 84.0, not estimable). For LADC, PTEN- (based on the pathologists' scores cut-off=0) was a negative prognostic factor for all endpoints: OS (p=0.004), RFS (p=0.002) and TTR (p=0.003) (data not shown). This negative effect of PTEN- is further demonstrated when dividing the H-score averages in 4 levels (≤5, >5 to ≤20, >20 to <100, ≥100) (Supplementary Figure 3A/B; $p<0.05$ for both pathologists and computer). In the adjusted Cox model depicted in the forest plot (Supplementary Figure 4A), the HR for PTEN- vs. + was 1.21 (95% CI: 1.01, 1.46; p=0.04) for the LADC and the pathologist threshold 0, indicating higher death risk by 21%. Similar results were obtained for pathologists' threshold 8, computer threshold 5 (Supplementary Figures 4B, 4C).

Finally, the KM-plotter database of 1145 independent NSCLC patients was interrogated for PTEN mRNA expression dichotomized at the median. PTEN expression in all NSCLC correlated with better OS (HR =0.49, $p<0.001$), matching our results (Supplementary Figure 5). Regarding histologic subtype, high PTEN expression in LADC, but not LSCC, was also associated with better OS (HR=0.41, $p<0.001$).

Comparison of pathologists versus computer scoring

The intra-class correlation coefficient (ICC) was calculated in order to evaluate the resembleness of H-scores of common cases among different centres participating in the EQA. The ICC of the EQA TMA PTEN stainings for pathologists was 0.72, while for the computer it was 0.66. The ICC of the EQA webbook for pathologists is 0.88, using the SP218 antibody. Moreover, the EQA TMA analysis showed a lower inter-center variability for computer-derived H-scores compared to the pathologists. The difference in calibration threshold settings between 3 internal Zurich pathologists and 10 ETOP pathologists was minimal (Figure 4A). Furthermore, the computer H-scores were on average 54% lower than pathologists (Figure 4B).

Regarding the ETOP cohort analysis, significant correlations were found for pathologists versus computer scores per individual core and per average of the 4 cores (all coefficients >0.55; $p<0.001$) (data not shown). Further analyses used the 4-core average. The frequency distributions of pathologists and computer scores are illustrated in Figure 4C using 5 H-score categories: 1) 0; 2) >0 to ≤5; 3) >5 to ≤20; 4) >20 to <100 and 5) ≥100. As in the EQA TMA analysis, the computerized methodology generally produced lower H-score values with the exception of H-scores close to 0. In the range close to 0, pathologists mostly scored exactly 0, whereas the computer mostly gave values ≤5. The opposite phenomenon was observed in the high expression range, where the pathologists scored higher than the computer.

Discussion

In this study, different IHC protocols were investigated in the multi-centre ETOP Lungscape cohort in order to measure PTEN protein immunoreactivity on tumor epithelia of formalin-fixed paraffin-embedded (FFPE) tissue blocks of resected chemo-naïve NSCLC patients. In our novel approach, pathologists calibrated the computerized image analysis of PTEN IHC. Thereby, pathologists' interpretations were directly compared with independent computer-

calculated H-scores on scanned slides obtained from pathologists' own intensity settings on an individual basis.

PTEN immunohistochemistry and FISH

Comprehensive PTEN IHC testing originally used 3 monoclonal antibodies (28H6, 10P03, 6H2.1) and one polyclonal on endometrial carcinoma. 6H2.1-derived immunoreactivity correlated best with *PTEN* gene alterations [30]. Recent studies in prostate, renal cell, breast, endometrial and vulvar carcinomas confirmed this finding [31-34]. Novel clones such as 138G6, SP218 and D4.3 were developed and tested on genetically defined PTEN +/- cell lines [35]. SP218 and 138G6 are RmAb with high affinities of kD 10^{-12} in contrast to MmAb with kD 10^{-9}. In this study the MmAb 6H2.1 although stained stronger than the two RmAbs. Based on EQA results, performance of 6H2.1 should rather be considered overstaining, as e.g. lung carcinomas with genomic PTEN- showed H-scores up to 63. The synthetic CTD peptide for SP218 and 138G6 is most likely the same, explaining the similar H-score values in the EQA.

Lung cancer may show complex genomic alterations with subclonal deletions. We used the *CEP10/PTEN* ratio to measure the global genomic *PTEN* status. Most PTEN- cases were close to hemizygous or homozygous loss, except the mucinous endometrium carcinoma with a value of 0.90. It is noteworthy that cells are ultrasensitive to subtle changes of *PTEN* dosage. Hypermorphic mice with a 20% reduction in mRNA *PTEN* levels developed already a spectrum of tumors [36].

Computerized analysis of semi-quantitative intensity marks

In our computerized pixel-based methodology chromogenic intensity levels were calibrated by averaged settings of surgical pathologists with >10 years of IHC experience. Such thresholds may serve as "intensity reference marks" to be put in the continuous variable of brown tone intensity. Evidently, the absolute position of the intensity marks 1, 2 or 3 on the 255 to 0 Lab scale would be dependent on the general staining intensity of a given antibody. Inhere, these marks were laid on the brown tone scale in a rather close and comparable

fashion by pathologists. It remains to be seen if the use of the entire brown tone intensity spectrum as continuous variable would add further value in comparison to the intensity mark approach. In our study, total cellular immunoreactivity was scored. A drawback of this approach is the lack of intracellular signal localization to either cytosol or nucleus.

Computerized analysis of automated quantitative analysis (AQUA)-based immunofluorescence allowed for measuring cytosolic PTEN- in NSCLC [37]. More recently, chromogenic PTEN IHC was measured in NSCLC in the red-green-blue (RGB) spectrum using Aperio ImageScope software and 6H2.1 antibody [34, 38]. A similar approach was used for scoring PTEN expression in larynx carcinoma [39]. Furthermore, its intracellular localization, intensity and frequency was assessed using 3D Histech Panoramic Viewer software and antibody clones 6H2.1, 138G6 and Ab6-28H6 [40].

Comparison of pathologists versus computer

ETOP pathologists were challenged with their own computerized intensity settings. They were asked to semi-quantitatively set intensity thresholds on a computer screen on representative images. Thereafter, they scored the EQA webbook on screen. These scores were compared with computerized analysis on the same images using their own intensity settings. More narrow data was obtained by the computer.

Scoring of brown tone intensity of immunohistochemistry by the human eye is assumed to be subjective. Agreement on score 0 and 3 is usually high. Scores 1 and 2 are more critical, but reasonable inter-observer kappa values of 0.7 or higher can be achieved if visual dictionaries are provided [33, 41]. The integration of various frequencies with different intensities is particularly challenging since the perception of immunoreactivity on tumor epithelia by the human eye will be influenced by the staining intensity of surrounding CAFs, a phenomenon called Chubb illusion. Chubb illusion is an error in visual perception in which the apparent contrast of an object varies substantially to most viewers depending on its relative contrast to the field on which it is displayed [42]. A high density of strongly PTEN positive CAFs therefore

tends to result in lower PTEN scores for the tumor epithelia when assessed by pathologists but not by the computer.

Regarding perception of stained area, over-scoring may occur due to the fact that strongly stained areas are preferentially recognized in terms of percent surface. Here lies the strength of objective pixel-based intensity measurements leading to approximately 50% lower but more homogenous scores. Inter-center variability was also lower when the computer was used, which is useful for multicentre studies. Pathologists downscaled small H-scores of up to 5 resulting from few brown pixels (computer), to the '0' value while brown intensities greater than 20 (computer) were overrated and given values often greater than 100.

Setting intensity threshold for PTEN-

Due to the absence of a consensus, different thresholds were considered: for pathologists 0 (a priori) and 8 (calculated) and for computer 5 (calculated). From a cell biological point of view it is conceivable that alternative cut-off values beside the a priori H-score 0 could be used for correlation with clinic-pathologic parameters. It is important but also challenging to set an optimal threshold in the lower range of H-scores. Computerized analysis is thereby of great help to distinguish small differences in brown tones. Admittedly, using the different thresholds 0, 5 or 8, no major differences in correlation between PTEN- and clinic-pathologic parameters were found.

Correlation with clinic-pathologic parameters

It has been shown that tumor cells are ultrasensitive to subtle PTEN dosage alterations, and a hypermorphic allele with a remaining 80% of wild-type activity increased tumor formation in mice [43]. Complete PTEN loss triggered cellular senescence, which protected against tumor initiation or progression [44]. Therefore, partial PTEN loss is advantageous for initial tumorigenesis, whereas complete loss promotes rapid tumor growth after senescence mechanisms are impaired in advanced tumors. Our data fit with this concept as larger tumors showed a stronger PTEN loss. Our results also confirm the dismal prognostic value of PTEN-

in early stage adenocarcinomas [45]. We observed a protein loss in 44% (pathologists, threshold 0) and 59% (computer, threshold 5) of NSCLC, respectively. The frequencies were 51%/65% for LSCC and 37%/54% for LADC. Our results thus match with other NSCLC studies where PTEN- occurred in 21% [20] or 59% [21] of LSCC, but 4% or 34% in LADC, respectively. These studies used the 138G6 antibody and PTEN- was defined as absence of any immunoreactivity.

Clinical significance

The main goal of medical expert systems is not to replace human experts, but rather to support them. In digital pathology, computerized image analysis using pixel-based measurement is of great help for many tasks- measurement of immunohistochemical brown staining across tumor surfaces. Thus, using the computer we add value to H-scores.

Here, we present a novel approach of pathologist's calibrated IHC analysis of PTEN protein. Pathologists thereby compare their interpretations with computer-calculated H-scores derived from their own intensity settings. Due to unbiased integration of pixel intensities in a given frame-of-interest, more homogenous H-scores are achieved. Such algorithms may improve reproducibility of biomarker detection regarding epitope validation in international clinical trials that stratify patients according to protein expression values.

Conclusion

In summary, calibration of immunoreactivity intensities by pathologists, following computerized H-score measurements, has the potential to improve reproducibility and homogeneity of biomarker detection regarding epitope validation in multi-centre studies. Such computerized measurements are able to deliver adequate data in the low expression range and to confirm pathologists' scorings regarding correlations with clinical-pathologic parameters, including survival.

Acknowledgments

This work was supported by grants from the Swiss Cancer League (reference number F-87701-31-01) and the Swiss National Science Foundation Systems X (reference number M-87704-01-02) to A.S. Roche/Genentech contracted with ETOP to design the Lungscape 002 PTEN study, and stain, score, collect, analyze and interpret the results. We also want to thank all technicians from all collaboration centers and Lungscape steering committee that helped to carry out the work with samples (Supplementary appendix).

Legends of figures and tables

Figure 1. PTEN immunohistochemistry. A) Comparison of the 3 antibodies using a PTEN-mucinous endometrium carcinoma from EQA. Cancer-associated fibroblasts (CAFs) are stained positive and serve as internal control. B) Representative H-scores for SP218 antibody indicated by pathologists. Genomic *PTEN* loss (CEP10/PTEN ratio 0.02) is demonstrated by FISH at 1000x magnification in LSCC from EQA; CEP10 (green), PTEN (orange), nuclei (blue). Small peripheral lymphocytes show both CEP10 (green) and PTEN (orange) signals, but tumor nuclei only CEP10. Upper (x-axis) and right (y-axis) border indicate distribution of fluorescent intensity across merged z-stack of 20 images.

Figure 2. Methodology of computerized analysis of PTEN immunoreactivity based on pathologists' calibrated pixel intensity measurement. In FIJI Lab color space, color threshold was adjusted, performing 3-fold background normalization and followed by averaged intensity calibration by pathologists. The threshold between brown signal intensity 1 and blue-grey color of hematoxylin/bluing reagent counterstaining with eventual faint brownish tone (intensity score 0) was averaged from 110 representative image frames laid on faintly stained glass, unstained tissue areas of a microfluidic tissue processor-based IHC and stainings without primary antibody. Thresholds for intensity 2 or 3 levels were averaged from 3 Zurich pathologists and used for H-score calculation. Note: For subsequent analysis of both EQA and ETOP cohort frames of size 100-500 x 100-500 pixels were laid on tumor epithelia only.

Figure 3. Setting optimal threshold in external quality assessment. A) Boxplots of H-scores for EQA webbook PTEN- and PTEN+ cases for the three antibodies scored by pathologists and computer. These results indicates that 6H2.1 antibody stains stronger. B)

ROC curves with optimal thresholds for PTEN- calculated from EQA (pathologists H-score value = 8, computer H-score value = 5) together with corresponding specificity/sensitivity (0.92/0.94) and (0.95/0.92) for predicting PTEN-.

Figure 4. Performance of pathologists versus computer. A) Brown tones intensity thresholds without background set by pathologists from Zurich (red triangle) and 10 ETOP centers (dots) on 3 image frames of 1 webbook case (PTEN+ LADC) stained with SP218. B) Comparison of H-score variation between pathologists and computer for 6 PTEN- (left, y-axis divided into 0 to 0.04 and 0 to 80 for 3 cases each) and 6 PTEN+ (right) webbook cases, stained centrally in Zurich. Computer H-scores for every center were calculated based on intensity thresholds set by the respective ETOP pathologists. Pathologists H-scores reflect own scoring of webbook in each ETOP center. C) Population tree comparing frequencies of H-score categories between pathologists and computer. H-scores of whole ETOP cohort from 97 TMAs stained with SP218 were divided in 5 categories 1) 0; 2) >0 to ≤5; 3) >5 to ≤20; 4) >20 to <100 and 5) ≥100.

Table 1. External quality assessment results. Summary of PTEN IHC H-scores averaged across ETOP centers for the 12 EQA cases ordered by histology. These cases were stained with SP218, 138G6 and 6H2.1 for central webbook and SP218 only for local TMA. All cases were stained with a PTEN FISH probe. EQA TMA (both computer and pathologists) and EQA webbook (pathologists, all 3 antibodies) data are based on 16 evaluations/scorings per case, while EQA webbook data evaluated by computer are based on 50 pictures per case for each antibody. FISH analysis was done by analyzing 100 cells for each case.

Table 2. Correlations of PTEN immunoreactivity using the SP218 antibody with clinic-pathologic parameters of Lungscape cohort patients. CT chemotherapy, RT radiotherapy.

Fischer's exact statistical test is used, unless indicated otherwise. (*) Categories 'Large cell' and 'Other' excluded, (†) Categories 'Large cell' and 'Other' combined, (‡) Mantel-Haenszel test.

List of supplemental data

Supplementary Video

Supplementary Figure 1

Supplementary Figure 2

Supplementary Figure 3

Supplementary Figure 4

Supplementary Figure 5

Supplementary Table 1

Supplementary Table 2

Supplementary Table 3

Supplementary Appendix

References

1. Baker SJ. PTEN enters the nuclear age. *Cell* 2007;128:25-28.

2. Perren A, Komminoth P, Saremaslani P, et al. Mutation and expression analyses reveal differential subcellular compartmentalization of PTEN in endocrine pancreatic tumors compared to normal islet cells. *Am J Pathol* 2000;157:1097-1103.

3. Yamamoto H, Shigematsu H, Nomura M, et al. PIK3CA mutations and copy number gains in human lung cancers. *Cancer Res* 2008;68:6913-6921.

4. Fumarola C, Bonelli MA, Petronini PG, et al. Targeting PI3K/AKT/mTOR pathway in non small cell lung cancer. *Biochem Pharmacol* 2014;90:197-207.

5. Leslie NR, Brunton VG. Cell biology. Where is PTEN? *Science* 2013;341:355-356.

6. Lotan TL, Wei W, Ludkovski O, et al. Analytic validation of a clinical-grade PTEN immunohistochemistry assay in prostate cancer by comparison with PTEN FISH. *Mod Pathol* 2016;29:904-14.

7. Edelman G, Rodon J, Lager J, et al. Phase I Trial of a Tablet Formulation of Pilaralisib, a Pan-Class I PI3K Inhibitor, in Patients with Advanced Solid Tumors. *Oncologist* 2018;23:401-e438.

8. Vestergaard HH, Christensen MR, Lassen UN. A systematic review of targeted agents for non-small cell lung cancer. *Acta Oncol* 2018;57:176-186.

9. Jamaspishvili T, Berman DM, Ross AE, et al. Clinical implications of PTEN loss in prostate cancer. *Nat Rev Urol* 2018;15:222-234.

10. George S, Miao D, Demetri GD, et al. Loss of PTEN Is Associated with Resistance to Anti-PD-1 Checkpoint Blockade Therapy in Metastatic Uterine Leiomyosarcoma. *Immunity* 2017;46:197-204.

11. Peng W, Chen JQ, Liu C, et al. Loss of PTEN Promotes Resistance to T Cell-Mediated Immunotherapy. *Cancer Discov* 2016;6:202-216.

12. Sharma MD, Shinde R, McGaha TL, et al. The PTEN pathway in Tregs is a critical driver of the suppressive tumor microenvironment. *Sci Adv* 2015;1:e1500845.

13. Leong JW, Schneider SE, Sullivan RP, et al. PTEN regulates natural killer cell trafficking in vivo. *Proc Natl Acad Sci USA* 2015;112:E700-709.

14. Xu C, Fillmore CM, Koyama S, et al. Loss of Lkb1 and Pten leads to lung squamous cell carcinoma with elevated PD-L1 expression. *Cancer Cell* 2014;25:590-604.

15. Kim HS, Lee JH, Nam SJ, et al. Association of PD-L1 Expression with Tumor-Infiltrating Immune Cells and Mutation Burden in High-Grade Neuroendocrine Carcinoma of the Lung. *J Thorac Oncol* 2018;13:636-648.

16. Fadejeva I, Olschewski H, Hrzenjak A. MicroRNAs as regulators of cisplatin-resistance in non-small cell lung carcinomas. *Oncotarget* 2017;8:115754-115773.

17. Lebok P, Kopperschmidt V, Kluth M, et al. Partial PTEN deletion is linked to poor prognosis in breast cancer. *BMC Cancer* 2015;15:963.

18. Ocana A, Vera-Badillo F, Al-Mubarak M, et al. Activation of the PI3K/mTOR/AKT pathway and survival in solid tumors: systematic review and meta-analysis. *PloS One* 2014;9:e95219.

19. Collaud S, Tischler V, Atanassoff A, et al. Lung neuroendocrine tumors: correlation of ubiquitinylation and sumoylation with nucleo-cytosolic partitioning of PTEN. *BMC Cancer* 2015;15:74.

20. Spoerke JM, O'Brien C, Huw L, et al. Phosphoinositide 3-kinase (PI3K) pathway alterations are associated with histologic subtypes and are predictive of sensitivity to PI3K inhibitors in lung cancer preclinical models. *Clin Cancer Res* 2012;18:6771-6783.

21. Yanagawa N, Leduc C, Kohler D, et al. Loss of phosphatase and tensin homolog protein expression is an independent poor prognostic marker in lung adenocarcinoma. *J Thorac Oncol* 2012;7:1513-1521.

22. Marsit CJ, Zheng S, Aldape K, et al. PTEN expression in non-small-cell lung cancer: evaluating its relation to tumor characteristics, allelic loss, and epigenetic alteration. *Hum Pathol* 2005;36:768-776.

23. Djordjevic B, Hennessy BT, Li J, et al. Clinical assessment of PTEN loss in endometrial carcinoma: immunohistochemistry outperforms gene sequencing. *Mod Pathol* 2012;25:699-708.

24. Vlietstra RJ, van Alewijk DC, Hermans KG, et al. Frequent inactivation of PTEN in prostate cancer cell lines and xenografts. *Cancer Res* 1998;58:2720-2723.

25. Peters S, Weder W, Dafni U, et al. Lungscape: resected non-small-cell lung cancer outcome by clinical and pathological parameters. *J Thorac Oncol* 2014;9:1675-1684.

26. Blackhall FH, Peters S, Bubendorf L, et al. Prevalence and Clinical Outcomes for Patients With ALK-Positive Resected Stage I to III Adenocarcinoma: Results From the European Thoracic Oncology Platform Lungscape Project. *J Clin Oncol* 2014;32:2780-2787.

27. Gyorffy B, Surowiak P, Budczies J, et al. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PLoS One* 2013;8:e82241.

28. Schindelin J, Rueden CT, Hiner MC, et al. The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol Reprod Dev* 2015;82:518-529.

29. Ciftlik AT, Lehr HA, Gijs MA. Microfluidic processor allows rapid HER2 immunohistochemistry of breast carcinomas and significantly reduces ambiguous (2+) read-outs. *Proc Natl Acad Sci USA* 2013;110:5363-536

30. Pallares J, Bussaglia E, Martinez-Guitarte JL, et al. Immunohistochemical analysis of PTEN in endometrial carcinoma: a tissue microarray study with a comparison of four commercial antibodies in correlation with molecular abnormalities. *Mod Pathol* 2005;18:719-727.

31. Garg K, Broaddus RR, Soslow RA, et al. Pathologic scoring of PTEN immunohistochemistry in endometrial carcinoma is highly reproducible. *Int J Gynecol Pathol* 2012;31:48-56.

32. Carvalho KC, Maia BM, Omae SV, et al. Best practice for PTEN gene and protein assessment in anatomic pathology. *Acta Histochem* 2013;116:25-31.

33. Maiques O, Santacana M, Valls J, et al. Optimal protocol for PTEN immunostaining; role of analytical and preanalytical variables in PTEN staining in normal and neoplastic endometrial, breast, and prostatic tissues. *Hum Pathol* 2014;45:522-532.

34. Lavorato-Rocha AM, Anjos LG, Cunha IW, et al. Immunohistochemical assessment of PTEN in vulvar cancer: best practices for tissue staining, evaluation, and clinical association. *Methods* 2015;77-78:20-24.

35. Lotan TL, Gurel B, Sutcliffe S, et al. PTEN protein loss by immunostaining: analytic validation and prognostic indicator for a high risk surgical cohort of prostate cancer patients. *Clin Cancer Res* 2011;17:6563-6573.

36. Alimonti A, Carracedo A, Clohessy JG, et al. Subtle variations in Pten dose determine cancer susceptibility. *Nat Genet* 2010;42:454-458.

37. Gustavson MD, Bourke-Martin B, Reilly D, et al. Standardization of HER2 immunohistochemistry in breast cancer by automated quantitative analysis. *Arch Pathol Lab Med* 2009;133:1413-1419.

38. Panagiotou I, Georgiannos SN, Tsiambas E, et al. Impact of HER2 and PTEN simultaneous deregulation in non-small cell lung carcinoma: correlation with biological behavior. *Asian Pac J Cancer Prev* 2012;13:6311-6318.

39. Mastronikolis NS, Tsiambas E, Papadas TA, et al. Deregulation of PTEN Expression in Laryngeal Squamous Cell Carcinoma Based on Tissue Microarray Digital Analysis. *Anticancer Res* 2017;37:5521-5524.

40. Agoston EI, Micsik T, Acs B, et al. In depth evaluation of the prognostic and predictive utility of PTEN immunohistochemistry in colorectal carcinomas: performance of three antibodies with emphasis on intracellular and intratumoral heterogeneity. *Diagn Pathol* 2016;11:61.

41. McCarty KS, Jr., Miller LS, Cox EB, et al. Estrogen receptor analyses. Correlation of biochemical and immunohistochemical methods using monoclonal antireceptor antibodies. *Arch Pathol Lab Med* 1985;109:716-721.

42. Chubb C, Sperling G, Solomon JA. Texture interactions determine perceived contrast. *Proc Natl Acad Sci USA* 1989;86:9631-9635.

43. Alimonti A. PTEN breast cancer susceptibility: a matter of dose. *Ecancermedicalscience* 2010;4:192.

44. Chen JJ, Lin YC, Yao PL, et al. Tumor-associated macrophages: the double-edged sword in cancer progression. *J Clin Oncol* 2005;23:953-964.

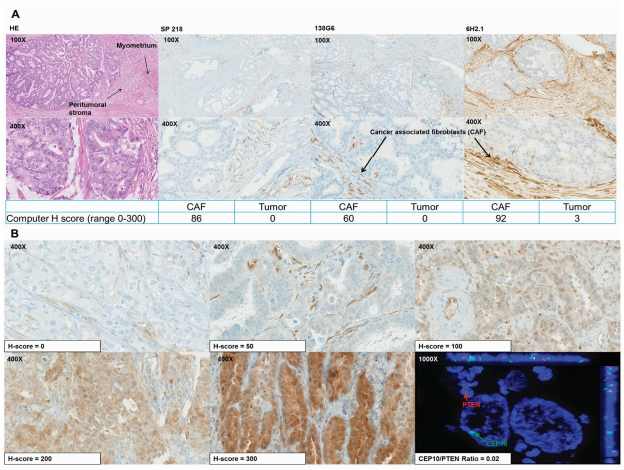
45. Gu J, Ou W, Huang L, et al. PTEN expression is associated with the outcome of lung cancer: evidence from a meta-analysis. *Minerva Med* 2016;107:342-51.

Table 1. External quality assessment results. Summary of PTEN IHC H-scores averaged across ETOP centers for the 12 EQA cases ordered by histology. These cases were stained with SP218, 138G6 and 6H2.1 for central webbook and SP218 only for local TMA. All cases were stained with a PTEN FISH probe. EQA TMA (both computer and pathologists) and EQA webbook (pathologists, all 3 antibodies) data are based on 16 evaluations/scorings per case, while EQA webbook data evaluated by computer are based on 50 pictures per case for each antibody. FISH analysis was done by analyzing 100 cells for each case.

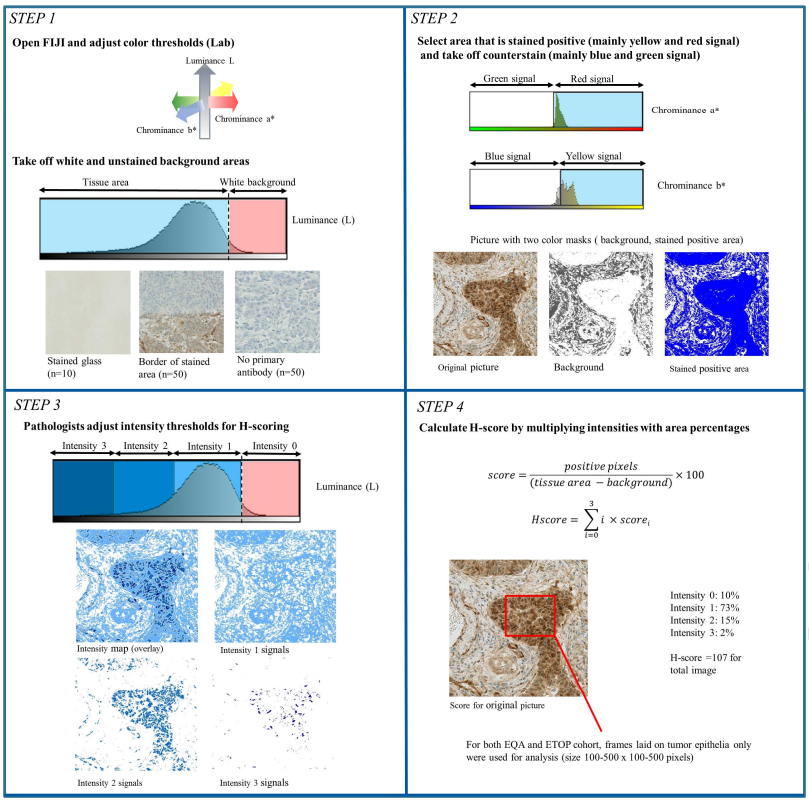
	Mean H-score								
	EQA TMA (local)		**EQA Webbook (central)**						**FISH**
	SP218		**SP218**		**138G6**		**6H2.1**		**CEP10/PTEN**
Diagnosis	**Pathologists**	**Computer**	**Pathologists**	**Computer**	**Pathologists**	**Computer**	**Pathologists**	**Computer**	
Endometrioid Endometrium Carcinoma									
PTEN-	17	4	13	1	4	0	23	59	0.56
Mucinous Endometrium Carcinoma									
PTEN-	1	1	0	0	0	0	4	2	0.90
Lung Adenocarcinoma									
PTEN+	97	53	96	23	103	28	277	187	0.99
PTEN-	1	0	1	0	1	0	17	5	0.60
PTEN-	1	1	7	0	1	0	53	63	0.58
Lung Squamous Cell Carcinoma									
PTEN+	190	77	209	130	178	62	245	121	1.09
PTEN-	0	0	1	0	7	0	18	6	0.02
Glioblastoma									
PTEN+	148	65	232	111	198	61	296	200	1.07
Prostate Carcinoma									
PTEN+	91	58	127	76	82	22	186	129	1.00
Solid Carcinoma Fallopian Tube									
PTEN+	68	30	149	73	98	10	248	123	1.11
Lung Carcinoma Cell Line H460									
PTEN+	133	54	131	36	228	123	265	143	1.04
Prostate Carcinoma Cell Line PC3									
PTEN-	17	8	0	0	0	0	25	1	0.59

Table 2. Correlations of PTEN immunoreactivity using the SP218 antibody with clinic-pathologic parameters of Lungscape cohort patients. CT chemotherapy, RT radiotherapy. Fischer's exact statistical test is used, unless indicated otherwise. (*) Categories 'Large cell' and 'Other' excluded, (†) Categories 'Large cell' and 'Other' combined, (‡) Mantel-Haenszel test.

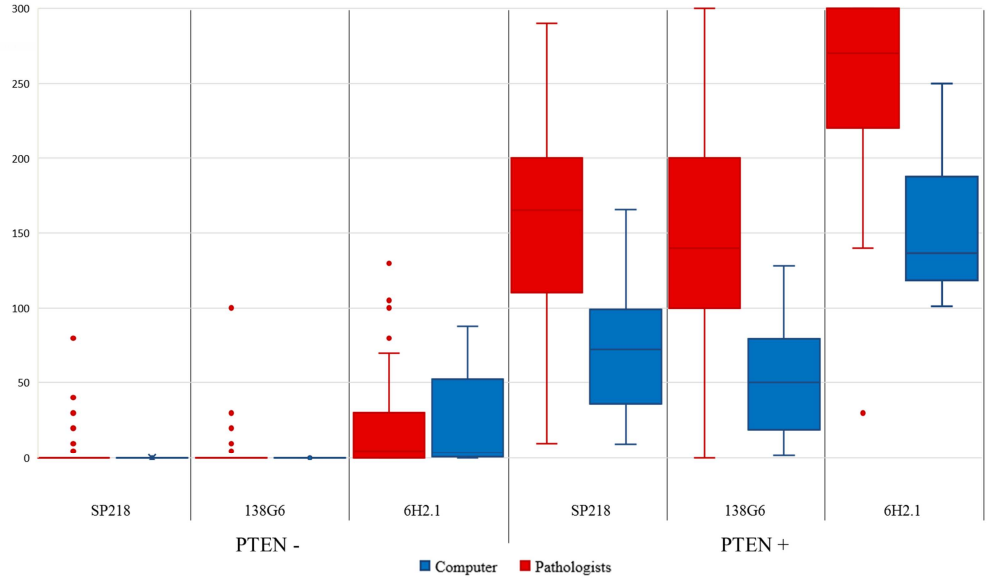
	Pathologists					Computer		
		Threshold: 0		Threshold: 8			Threshold: 5	
Characteristic	**n=2245**	**PTEN- n=981 (43.7%)**	**p-value**	**PTEN- n=1097 (48.9%)**	**p-value**	**n=2208**	**PTEN- n=1304 (59.1%)**	**p-value**
Age at surgery - n (%)								
<60	634	274 (43.2)	0.91	313 (49.4)	0.90	622	377 (60.6)	0.65
60-70	804	350 (43.5)		388 (48.3)		794	466 (58.7)	
>70	806	357 (44.3)		396 (49.1)		791	461 (58.3)	
Gender - n (%)								
Male	1479	698 (47.2)	**<0.001**	770 (52.1)	**<0.001**	1450	866 (59.7)	0.39
Female	766	283 (36.9)		327 (42.7)		758	438 (57.8)	
Smoking history - n (%)								
Former	1210	554 (45.8)	**<0.001**	615 (50.8)	**<0.001**	1191	702 (58.9)	**<0.001**
Current	707	337 (47.7)		374 (52.9)		696	438 (62.9)	
Never	237	58 (24.5)		72 (30.4)		230	106 (46.1)	
Adjuvant CT - n (%)								
Yes	495	258 (52.1)	**<0.001**	288 (58.2)	**<0.001**	486	311 (64.0)	**0.014**
No	1457	620 (42.6)		694 (47.6)		1431	824 (57.6)	
Adjuvant RT - n (%)								
Yes	111	67 (60.4)	**<0.001**	77 (69.4)	**<0.001**	110	72 (65.5)	0.16
No	1827	798 (43.7)		890 (48.7)		1793	1047 (58.4)	
Histology - n (%)								
Adenocarcinoma	1093	409 (37.4)	**<0.001**	456 (41.7)	**<0.001**	1073	581 (54.1)	**<0.001**
Squamous cell	983	503 (51.2)	**<0.001***	563 (57.3)	**<0.001***	970	628 (64.7)	**<0.001***
Large cell	98	46 (46.9)	**<0.001†**	48 (49.0)	**<0.001†**	95	60 (63.2)	**<0.001†**
Other	71	23 (32.4)		30 (42.3)		70	35 (50.0)	
Stage - n (%)								
I	1083	438 (40.4)	**0.002‡**	495 (45.7)	**0.0021‡**	1068	588 (40.4)	**<0.001**
II	653	297 (45.5)		329 (50.4)		640	391 (45.4)	**<0.001‡**
III	509	246 (48.3)		273 (53.6)		500	325 (65.0)	
Tumor size - n (%)								
≤4 cm	1394	541 (38.8)	**<0.001**	610 (43.8)	**<0.001**	1371	748 (54.6)	**<0.001**
>4 cm	849	439 (51.7)		485 (57.1)		835	556 (66.6)	



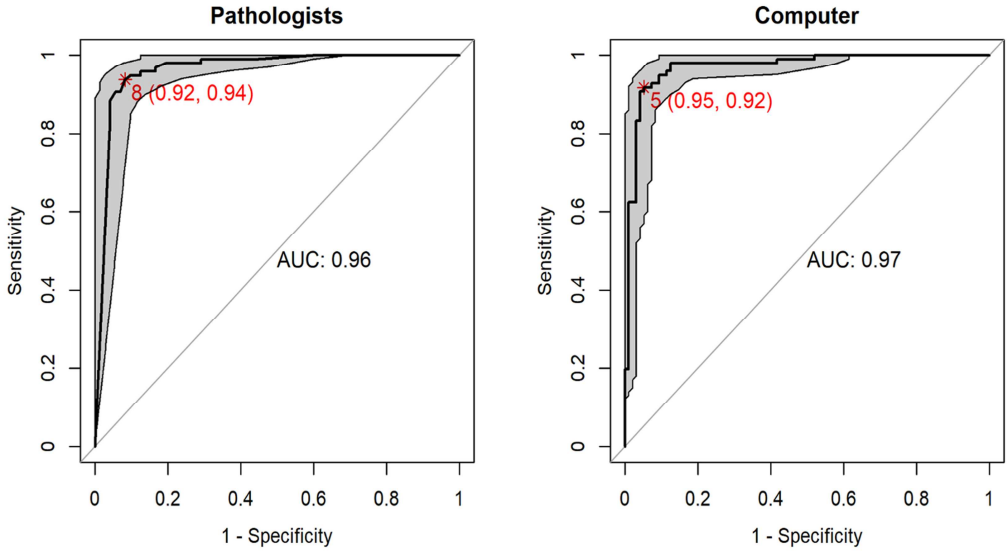
	CAF	Tumor	CAF	Tumor	CAF	Tumor
Computer H score (range 0-300)	86	0	60	0	92	3



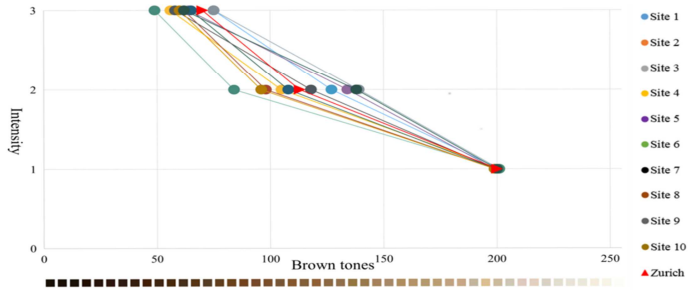
A.


300
250
200
150
100
50
0
SP218
138G6
6H2.1
SP218
138G6
6H2.1
PTEN -
PTEN +
Computer
Pathologists

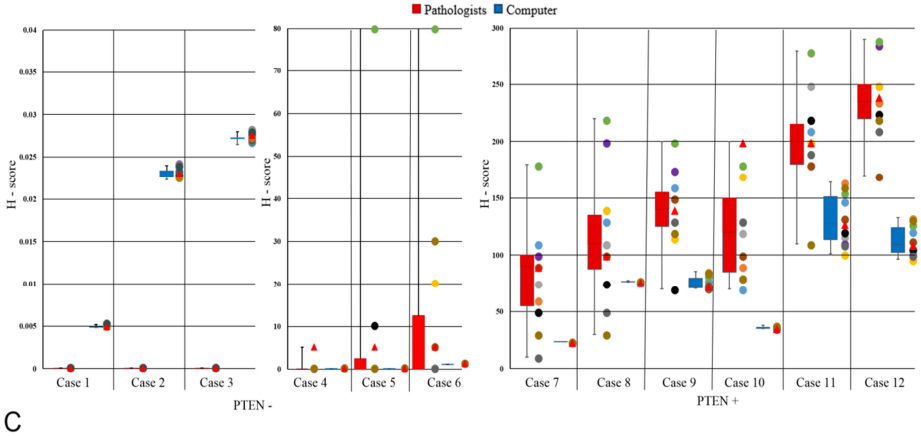
B.


Pathologists
8 (0.92, 0.94)
AUC: 0.96
Sensitivity
1 - Specificity
Computer
5 (0.95, 0.92)
AUC: 0.97
Sensitivity
1 - Specificity

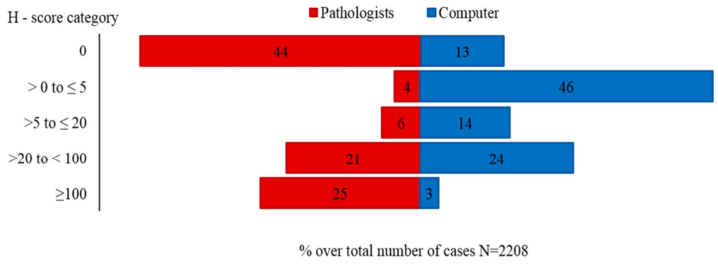
A


Intensity
Brown tones
Site 1
Site 2
Site 3
Site 4
Site 5
Site 6
Site 7
Site 8
Site 9
Site 10
Zurich

B


Pathologists
Computer
H - score
Case 1
Case 2
Case 3
Case 4
Case 5
Case 6
Case 7
Case 8
Case 9
Case 10
Case 11
Case 12
PTEN -
PTEN +

C


H - score category
Pathologists
Computer
0
> 0 to ≤ 5
>5 to ≤ 20
>20 to < 100
≥100
44
13
4
46
6
14
21
24
25
3
% over total number of cases N=2208